PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : H01J 49/04

A1

(11) International Publication Number: WO 98/09314

(43) International Publication Date: 5 March 1998 (05.03.98)

(21) International Application Number: PCT/GB97/02278

(22) International Filing Date: 26 August 1997 (26.08.97)

(30) Priority Data:
9617852.0 27 August 1996 (27.08.96) GB

(71) Applicant *(for all designated States except US)*: THE MANCHESTER METROPOLITAN UNIVERSITY [GB/GB]; All Saints Building, All Saints, Manchester M15 6BH (GB).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: CLAYDON, Martin, Arthur [GB/GB]; Needham Hall, 206 Palatine Road, Didsbury, Manchester M20 2TU (GB). DAVEY, Simon, Nicholas [GB/GB]; 16 Percival Street, Accrington, Lancs. BB5 0BE (GB). EDWARDS JONES, Valerie [GB/GB]; 24 Albury Drive, Norden, Rochdale, Lancs. QL12 7SX (GB). GORDON, Derek, Boyd [GB/GB]; 12 Knowl Close, Ramsbottom, Bury, Lancs. BL0 9YY (GB).

(74) Agents: EVANS, David, Charles et al.; F.J. Cleveland & Company, 40/43 Chancery Lane, London WC2A 1JQ (GB).

(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, HU, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).

Published
With international search report.

(54) Title: MICROORGANISM IDENTIFICATION

(57) Abstract

The invention relates to a method and apparatus for characterising microorganisms. The invention uses matrix assisted laser desorption ionisation time of flight mass spectrometry (MALDI-TOF-MS) spectral data for a range of known microorganisms. This spectral data is included in a database for spectral data of each of known microorganisms. A sample of an unidentified microorganism is prepared and then subject to analysis by MALDI-TOF-MS techniques to produce the spectral data of the unknown organism. This data is then compared; using suitable comparison means with the spectral data in the database. Similarity indicates the genus while identity indicates the specific microorganism. The apparatus for this analysis has the advantage that once set up and calibrated, it is relatively simple to use.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece			TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	NZ	New Zealand		
CM	Cameroon			PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

Micro-organism Identification

The present invention concerns the identification of micro-organisms, and more particularly to the differentiation between strains of bacteria.

The accurate identification of micro-organisms is time-consuming and complicated. As antibiotic-resistant strains of bacteria emerge there is a burgeoning need for fast and accurate identification of strains so that appropriate methods to combat them can be derived. Differentiation between strains of a bacterium is a very difficult process. Attempts have been made to distinguish between strains by the use of typing methodologies based upon phenotypic properties. The anti-biogram method attempts to distinguish differing strains by means of the variation in antibiotic resistance of each strain. Phage-typing distinguishes bacteria by infection with a virus.

However, while such methods can group bacteria according to a range of phenotypic properties, they are not very effective at distinguishing between genera or strains of bacteria which have almost identical genetic constitution. Bio-typing has been used to try to provide the required differentiation. Bio-typing investigates the differences in, for example, toxin production and trytophan dependence between different strains. Even this

method frequently fails to identify strains, even where the strains are derived from different patients and where a nosocomial incident (cross-infection) has occurred.

Tracing the source(s) of nosocomial incidents is vital in places, such as hospitals where such incidents can have severe health implications. Correct identification of the strain of organism is vital to the prevention of further infection and to the treatment of those infected. This is particularly important with the recent emergence of Methycillin Resistant Staphylococcus Aureus (MRSA), many strains of which cannot be phage-typed.

The analysis of mycobacteria present particular problems. The pathogenic forms, such as Mycobacterium tuberculosis, are typically slow to grow and accurate identification by traditional methods can take many weeks. However, early identification of tuberculosis is vital in the management of the infection in patients. Chromatographic analysis of derivatised mycolic acids and cell wall lipids is the current method for identification of mycobacteria. However such a method requires relatively large amounts of culture for the extraction and derivatisation steps.

Identification by mass spectrometry combined with pyrolysis (Py-MS), gas chromatography (GC-MS) and fast

atom bombardment (FAB-MS) has been applied to many areas of microbiology. However these methods require complex sample preparation and extraction procedures prior to analysis. In addition they require experienced specialist instrument operators which renders the microbiologist dependent upon external laboratories where such facilities can be maintained.

The use of Matrix-Assisted Laser Desorption Ionisation Mass Spectrometry (MALDI-TOF-MS) for characterisation of biopolymers has been reported by Hillenkamp et al in 1991 Anal. Chem. 63, 1193A-1202A.

Cotter has reported time of flight mass spectrometry for the structural analysis of biological molecules in 1992 Anal. Chem. 64, 1027A-1039A.

There is a need for a rapid, accurate method of differentiating between organisms, which method is capable of differentiating between different strains where the differences are normally difficult to characterise.

The present applicants have found that MALDI-TOF-MS spectral data for micro-organisms is spectral data which is dependent on the molecular structure of the micro-organisms concerned. In effect, therefore, such data is

a 'finger print' for the organism concerned. As a micro-organism transmogrifies, its MALDI-TOF-MS spectral data will change to reflect the change in structure/composition of the organism.

According to one aspect of the present invention, therefore, there is provided a method of characterising micro-organisms characterised by:-

providing a database of matrix assisted laser desorption ionisation time of flight mass spectrometry (MALDI-TOF-MS) spectral data for a range of known micro-organisms.

preparing a sample of unidentified micro-organisms and obtaining spectral data thereof by MALDI-TOF-MS

and comparing using suitable comparison means the spectral data so obtained with database to identify a known micro-organism having the same or similar spectral data.

In another aspect of the present invention there is provided apparatus for the screening of micro-organisms characterised in that the apparatus comprises:-

spectrometric means for producing MALDI-TOF-MS

spectral data of a sample organism .

database means containing MALDI-TOF-MS spectral data for a range of micro-organisms.

The invention enables the identification of micro-organisms quickly and without the labour intensive procedures hitherto employed.

MALDI-TOF-MS is a high performance mass spectrometry tool, which may be operated by the non-specialist. Furthermore, the equipment required to perform the analysis is relatively inexpensive. The method and/or apparatus of the invention uses dried, solid samples which are easy to handle and store, and which are capable of analysis in accordance with the invention over a wide mass range.

Prior to analysis, the micro-organism may be cultured by conventional means, for example, on a nutrient agar plate.

The spectrometer samples may be prepared by taking micro-organism cells e.g. from a culture medium and applying them directly to a sample plate and a matrix added. Alternatively the cells may be admixed with the matrix prior to application so that cellular integrity may be

retained.

The dried matrix-sample mixture is bombarded with laser energy to create gas phase ionic species which are then pulsed into a flight tube. Both positive and negative ions may be generated and identified.

The species may be identified by their mass to charge ratio (m/z), which can be identified with an accuracy of better than 0.1%. The m/z value of each spectral peak may be identified from the centroid of the peak corresponding to the to the average molecular mass of the particular $M+H^+$ ion (in the positive ion mode).

In another aspect of the invention, the spectral data of the unidentified micro-organism is derived from a plurality of laser shots of the sample. Preferably the position of the laser energy impinging on the sample is varied, between shots, over the sample spot. In this way the effect of heterogeneities in the sample may be reduced by the summing of spectra from each shot and/or by the exclusion of data statistically deviant from reasonable values. This multiple shot technique may assist accurate, reproducible spectral data to be produced. Preferably spectra are taken from each and every tracking position over the sample area.

A linear analyzer may be used to enhance the sensitivity of the data where limited amounts of sample are available.

The inventors of the present application have discovered that MALDI-TOF-MS produces unique spectra for each micro-organism allowing identification and differentiation between different genera of micro-organisms and even within different strains.

The present invention is also useful for the analysis of mycobacteria, since the MALDI-TOF-MS method requires very small amounts of material, can be completed within minutes, and requires no derivatisation step.

In another aspect of the present invention each spectra compiled in the database is recorded as a graphical representation of a plot of intensity against mass to charge ratio.

The spectral data may be recorded in graphic, numerical or electronic format, in digital or analogue form. Preferably each spectra is recorded on the storage medium of a computer or computer network. The storage medium may be magnetic, such as floppy disk, tape or hard disk. Alternatively the storage medium may be optical, such as CD-ROM or laser-disc. in another embodiment the storage

medium may be ROM microchips.

Preferably each spectra is stored as a series of mass to charge ratios each corresponding to a centroid of the peaks of each spectrum. By storing the mass to charge ratios alone, the minimum data required to characterize and/or identify the micro-oranisms is stored.

To aid the search and identification process, the data may be arranged in groups of data corresponding to the genus of each micro-organism, with sub-divisions corresponding to the strain of micro-organism.

Where a spectrum analyzed is not capable of identification, means may be provided to add the spectrum to the database for future reference.

Where the database is stored in the storage medium of a computer, software may be provided to carry out the screening of the unidentified spectral data.

The software may carry out the steps of acquiring and storing the unidentified micro-organism's spectral data by manual input or, preferably directly from the spectrometer, comparing the unidentified spectral data with the pre-characterised spectral data in the database and indicating where a match is found, or otherwise.

In one embodiment, the database may include spectral data grouped according to the genus of each micro-organism and the software may be carry out a search strategy which involves comparing spectral data first to identify the genus of the unidentified bacteria and then to identify the strain of bacteria. In this way the search process is optimized to improve efficiency where a large range of spectral data is stored in the database.

The software may be capable of allowing for calibration errors in the unidentified spectral data, so that the trends (e.g. the difference in m/z value between adjacent peaks in the data) may be screened with the corresponding data in the database rather than the absolute values.

The computer may be incorporated into the MALDI-TOF-MS Spectrometer so that data is transferred directly from the detector to the computer for analysis.

Alternatively the computer may be spaced apart from the spectrometer and linked by remote means.

The present invention also includes a database of micro-organisms comprising spectral data of each micro-organism obtained by MALDI-TOF-MS analysis for use in the method, or in conjunction with the apparatus, of the invention.

Following is a description by way of example only and with reference to the figures of the drawings of a method of carrying the present invention into effect.

Figure 1 is a graphical representation of partial spectra obtained from samples including SA, CF and ECC2.

Figure 2 is a graphical representation of a partial, positive ion, mass spectrum of Mycobacterium smegmatis, in the m/z range = 500-1500 Da. Each of the listed peaks A-K represent a number of possible isobaric species of mycoserates. The relative peak heights are related to the number of isobaric species present at each molecular weight. The peak letter, molecular weight, and one of the possible isobaric structures are listed as follows:

A:1156. $I_{19}II_{18}III_2IV_2V_{14}$ B:1170.$I_{17}II_{17}III_3IV_2V_{14}$ C:1184.$I_{16}II_{18}III_4IV_2V_{14}$

D:~~1198.I_{19}~~$II_{18}III_3IV_2V_{14}$ E:1212.$I_{17}II_{17}III_4IV_2V_{14}$ F:1226.$I_{16}II_{18}III_5IV_2V_{14}$

G:1240.$I_{19}II_{18}III_4IV_2V_{14}$ H:1254.$I_{17}II_{17}III_5IV_2V_{14}$ I:1268.$I_{17}II_{17}III_5IV_2V_{15}$

J:1282.$I_{18}II_{18}III_5IV_2V_{14}$ K:1296.$I_{18}II_{18}III_5IV_2V_{15}$

The molecular weights quoted above are for the mono-isotropic, molecular ion (M^+).

In the following example ten micro-organisms were analyzed to produce the MALDI-TOF-MS spectra. The organisms were:-

Micro-organism	
ESCHERICHIA COLI NCIMB8581	ECN
ESCHERICHIA COLI CJ53-R1NCTC50001	ECC1
ESCHERICHIA COLI CJ53-NCTC50167	ECC2
ESCHERICHIA COLI K12λEMG2	ECK
STAPHYLOCOCCUS SAPROPHYTICUS	SS
STAPHYLOCOCCUS AUREUS	SA
STAPHYLOCOCCUS EPIDERMIDIS	SE
CITROBACTER FREUNDII CF	CF
KLEBSIELLA AEROGENES	KA
MYCOBACTERIUM SMEGMATIS	MS

The ten micro-organisms were each sub-cultured in duplicate onto separate nutrient agar plates. Each pair was incubated at 37°C, for three and six days respectively.

A single, visible colony of cells was regarded as sufficient material for analysis and was removed from the agar plate with a sterile loop. All samples were analyzed

in a matrix of α-cyano-4-hydroxycinnamic acid (α-CHCA). The organisms were emulsified in 50-100 ul of the matrix, dissolved in a mixture of water, ethanol and acetonitrile (in the ratio 1:1:1).

A 1 ul aliquot was then applied to the sample target for the laser and allowed to dry before analysis. The samples were then analyzed by MALDI on a KRATOS™ Kompact MALDI III time of flight mass spectrometer. A nitrogen laser giving a 337 nm output of 3 ns pulse width was use to ionize the species. The laser fluence was set just above the threshold for ion production. The mass spectrometer was used in the positive ion detection mode using an acceleration voltage of +20 KV. Samples were analyzed by using both the linear and reflectron analyzers. The detector used consisted of a discrete dynode combine with an electron multiplier.

A spectrum was produced for each pair of each micro-organism. Figure 1 shows the spectra of SA, CF and ECC2. There is a clear difference in the distribution of the peaks corresponding to ionic species from each micro-organism. The spectra produced constitute a unique "fingerprint" which allows identification of micro-organisms inter-species and inter-strain. The similarities in the features of each spectra allow the spectra to be grouped according to genus of micro-

organism.

Figure 2 shows the positive ion spectrum of Mycobacterium smegmatis which shows an interesting pattern of peaks between 1150-1300 Da. Previous analyses of cell wall extracts from this organism suggest that these peaks are either mycolic acids or related lipids within the cell wall. The data probably relates to dimycocerosates of the phthiocerol family. These are non-covalently bound and are probably extracted into the matrix during the admixture process. The proposed structure of the mycoserates are shown in the legend of figure 2.

Table 1 lists subsets of m/z values for all the micro-organisms examined.

Unique patterns of m/z values from desorbed cell wall moieties obtained by MALDI-TOF-MS spectrometry									
E.Coli				Other Genera			Staphylococci		
ECN	ECC1	ECC2	ECK	CF	KA	MS	SS	SA	SE
							602	**602**	**602**
				617	617				
						618	**618**	**618**	**618**
							624	**624**	**624**
							642	**642**	**642**
						565	**656**	**656**	**656**
								715	715
					750				
				751					
							756		756
							800	**800**	**800**
							844	**844**	**844**
							888		
									989

						1021		1036	
		1058	1058						
		1093							
1102	**1102**	1102	**1102**	**1102**	**1102**				
						1158			
						1172			
						1186			
		1187	1187						
						1200			
						1214			
		1222							
						1228			
1231	**1231**	1231	**1231**	**1231**	**1231**				
						1242			
						1256			
						1270			
						1284			
						1298			
					1314				
1316	**1316**	1316	**1316**						
	1351	1351							
1360	**1360**	1360	**1360**	**1360**	**1360**				
		1445		1445					
		1480		1480					
1489	**1489**	1489	**1489**	**1489**	**1489**				
							1532		1532
				1543	1534				
		1574							
				1575					
1618	**1618**	1618	**1618**	**1618**	**1618**				
		1747		1747	1747				
1852			1852		1852				
				1876					

The m/z ratios of the spectral peaks of the four cultures of E.COLI bacteria are shown in the first column of table 1 for the m/z range 1058-1852 Da. The figure in bold type are peaks present in all the tested strains of the genus E.COLI (1102, 1231, 1316, 1360, 1489 and 1618). Thus the genus may be identified by the presence of these peaks.

The strain may be identified by the absence or presence of further peaks.

For example ECN has a peak at 1852 which differentiates from ECC1 and ECC2. ECK also has a peak at 1852, but may be differentiated by the additional peaks at 1058 and 1187.

Different species which are closely related to E.COLI are CF and KA. These have peaks 1102, 1231, 1360, 1489, 1618 in common with E.Coli, but may be differentiated therefrom by the absence of the peak at 1316 which is characteristic of E.Coli. CF and KA may be distinguished by the absence or presence of the peaks at 750, 751, 1314, 1445, 1480, 1575, 1582 and 1876.

The spectra produced, expressed in terms of mass to charge ratio of peaks in the spectra, provide a "fingerprint" for each micro-organism. The fingerprint is sensitive to both inter-genus differences, and inter-strain differences. A database of these fingerprints provides a method of rapidly screening samples to identify the strain of micro-organism present. This has particular application in the field of medical microbiology, but will find application in many other fields where the precise and rapid identification of micro-organisms is required.